Explaining Credibility in News Articles using Cross-Referencing

Dimitrios Bountouridis, Mónica Marrero, Nava Tintarev, Claudia Hauff
Delft University of Technology, The Netherlands
{d.bountouridis,m.marrero,n.tintarev,c.hauff}@tudelft.nl

ABSTRACT

The proliferation of online news sources has placed the issue of credibility at the center of public and scholarly attention alike. Without an authoritative entity that can vouch and sufficiently explain the quality of a piece of information appearing in news articles, readers become skeptical. At the same time, computational solutions are typically founded on different and possibly narrow interpretations of the complex concept of credibility. As such, while significant progress has been made, computational efforts have yet to propose a widely accepted solution. This paper proposes an interactive interface alternative to the existing algorithmic solutions: an additional information layer that is applied to an article's original textual contents. By contrasting heterogeneous articles of the same story, i.e., articles from different news outlets, our proposed approach reveals those pieces of information that are cross-referenced and thus—we argue—more likely to be credible. A demo of the tool is available at http://fairnews.ewi.tudelft.nl/InCredible/, the code is open-sourced at https://github.com/dbountouridis/InCredible.

KEYWORDS

credibility; explainability; news community; HCI, graph algorithms; computational journalism

ACM Reference format:
Dimitrios Bountouridis, Mónica Marrero, Nava Tintarev, Claudia Hauff. 2018. Explaining Credibility in News Articles using Cross-Referencing. In *Proceedings of SIGIR 2018 Workshop on ExplainAble Recommendation and Search, Ann Arbor, Michigan, USA, July 12, 2018 (EARS'18),* 5 pages.

1 INTRODUCTION

The World Wide Web has changed not only the way readers consume information, but also the way that information is created and disseminated. Regarding online news articles in particular, easy access to the Internet has allowed non-traditional entities e.g., individuals with limited journalistic experience, to spread information easily and with low cost. As a result, users have now access to news stories from a very diverse range of sources, i.e., news outlets, blogs, social media and so on. The amount of information available from various heterogeneous news sources places at the center of attention the issue of *credibility* [12]: readers are now suspicious with regards to the quality, believability and trustworthiness of the information provided to them [4]. We argue that this behavior relates to a lack of a trusted authority that can both vouch for the credibility of a piece of information and also explain how this credibility estimate came about.

Computational solutions that are designed to provide such estimates typically require formalized definitions of the concepts they aim to model; and credibility is a complex and multifaceted one [12]. Wathen and Burkell [15] identify a range of factors that influence the credibility of online information, such as trustworthiness, plausibility or similarity to the reader's beliefs. Cazalens et al. [4] investigated a number of fundamental issues in computational *factuality*-estimation, including the lack of a proper theoretical foundation and the lack of transparency in most state-of-the-art solutions. At the same time, factuality-estimation is problematic due to the subjective nature of *bias*; the question of factuality may appear only when an article's world-view is contrasted to the reader's [4, 15].

A number of prior works have proposed paradigms that can potentially overcome the factuality- and bias-related limitations. For instance, both Yin et al. [16] and Vydiswaran et al. [14] base their works on the idea that facts are persistent and that a piece of information does not need to be fact-checked to be considered credible, but rather is likely to be credible if it is cross-referenced i.e., it appears in multiple sources—a notion we also exploit in our work. Regarding the bias-related limitations, a number of works [7, 8, 13] and collective journalism projects[1] focus their efforts on providing readers with the "necessary" information to form their own bias conclusions about the information they receive. Consequently, the credibility judgment is only aided by the system and not replacing the reader's judgment [5]. At the same time, the main responsibility of those systems is to gather the necessary information and effectively communicate it to the reader. This implies a focus on *transparency* and *explainability*, i.e., the justification of why certain pieces of information are more likely to be credible than others.

This paper proposes a framework for aiding the credibility judgment in news articles based on the two aforementioned ideas of (i) cross-referenced information, and (ii) its proper, explainable communication to the reader. We hypothesize that well-communicated, cross-referenced information is more likely to be perceived as credible by the readers. The contributions of the paper are twofold:

- In Section 3 we propose an interactive interface that allows readers to identify which pieces of information in a news article are *corroborated* (i.e., the same piece of information exists in other articles on the same story), as well as those that are *omitted* (i.e., the piece of information exists in other articles on the same story but not the current one).
- In Section 4 we introduce a computational method to automatically determine corroborated and omitted pieces of information; in Section 5 we describe the corroboration/omission patterns we observe across different news outlets.

2 RELATED WORK

This section presents a number of works related to the interface component and computational backbone of our approach. Regarding the first component, the idea of exploiting knowledge from

EARS'18, July 12, 2018, Ann Arbor, Michigan, USA
2018. ACM ISBN .

[1] e.g., https://firstdraftnews.org/project/crosscheck/, retrieved May 2018

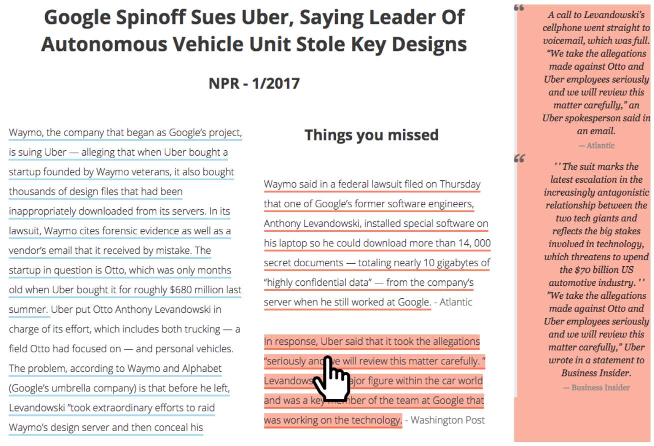

Figure 1: Overview of our proposed interactive information layer design: corroborated (blue underline) and omitted information (orange underline). Here, hovering over an underlined piece of information reveals information from other news sources that are omitted in the original story.

different sources has been explored in the past. For instance, a number of works (both academic and nonacademic) focus on the ideal of deliberative democracy, i.e., users can make better decisions only when exposed to diverse viewpoints [2]. `rbutr`[2] is an online initiative that aims to link web pages of contradicting content as manually annotated by its users. `DiversiNews` [7] provides a news aggregator system with a user interface that allows article filtering based on location, sentiment, and other features. `NewsCube` [13] aims at mitigating bias by providing readers with a selection of articles encapsulating all possible viewpoints. Their automated approach to locating all viewpoints is based on unsupervised clustering applied to the textual content. Similarly, `NewsBird` [8] is an aggregator framework that aims for the exploration of both common and different information. The News Landscape (NELA) toolkit [10] provides content-based analysis tools that make use of lexicon-based reliability and bias features among others. The online NELA[3] allows users to rank articles from various sources with respect to their reliability and bias estimates. It also provides a visual comparison of news sources using a variety of features. We note, that all mentioned aggregator interface approaches provide information at the article-level and not at the more fine-grained sentence level—in our work we tackle the latter.

Algorithmic solutions at the sentence-level focus on the notion of "trust". Systems such as [14] aim to estimate a numeric trust value for certain claims by assuming that reliable sources provide reliable claims and vice versa. Such systems are typically based on an iterative off-line computation (trust propagation) of both source- and claim-trust until their convergence. They are limited by two factors: the static nature of the estimate (undesirable in constantly updated collections of news articles) and the black-box nature of the underlying algorithms, which are not easily explainable to users.

[2] http://rbutr.com/, retrieved May 2018
[3] http://nelatoolkit.science/, retrieved May 2018

3 USER INTERFACE

Our main assumption is that well-communicated and cross-referenced information is more likely to be credible, in line with [14, 16]. As such, we now describe our user interface for effectively communicating cross-referenced information to the reader, while Section 4 describes our pipeline for locating that information.

The original content of the reader's selected online article remains unchanged; thus it can also be read in the traditional fashion. Our proposed and implemented interactive user interface (see Figure 1) acts as an additional information layer applied on top of this article, indicating whether certain pieces of information are corroborated or omitted:

Corroborated. Our design first highlights those pieces of information that are *corroborated*, meaning cross-referenced by other sources including the current article. Those pieces of information are underlined in blue, for example "President Trump visited France". Corroborated pieces of information in practice indicate that their content is likely to be credible since we use cross-referencing as a proxy of credibility. When hovered over, corroborated information reveal those pieces of information from other sources that are semantically similar; for example, "The president of the United States flew to France last weekend. - CNN" and "Donald Trump visited Europe for the second time last week. - NPR". This scheme explains to the readers why the current information is considered credible since it allows them to immediately contrast the content from different sources and make their personal credibility judgment.

Omitted. Under the heading "Things you might have missed", our proposed design additionally provides those pieces of information that are *omitted*, meaning cross-referenced by a number of other sources not including the current one. These pieces of information are underlined in orange, for example, "President Trump met with the French prime minister - CNN". Omitted information can be also hovered over to reveal the cross-referenced content from other sources. Such information is crucial not only for having a broader, more accurate picture of the news story but also for understanding whether the current source has purposefully (or not) decided to conceal likely-credible information from the reader.

It should be noted that under our proposed scheme, the cross-referenced pieces of information are the same for all articles of the same story, but they may appear as corroborated or omitted depending on the current article that the reader is viewing.

4 COMPUTATIONAL CROSS-REFERENCING

While Section 3 focused on communicating cross-referenced pieces of information (POIs from now on) via a user interface, this section describes our computational approach to automatically locate such information. Given a set $\mathcal{A}_s$ of articles pertaining to the same story s (existing topic detection and tracking approaches can be employed to determine this set for each story [1]), our pipeline consists of three stages (see Figure 2): (1) text segmentation, (2) POI similarity, and (3) graph generation and clique location.

Text segmentation. First, each article (or document) is split into POIs, i.e., meaningful chunks of text of one or more sentences that contain an argument, a quote, a fact and so on. Segmenting the text of a document into POIs is crucial for our pipeline. Our

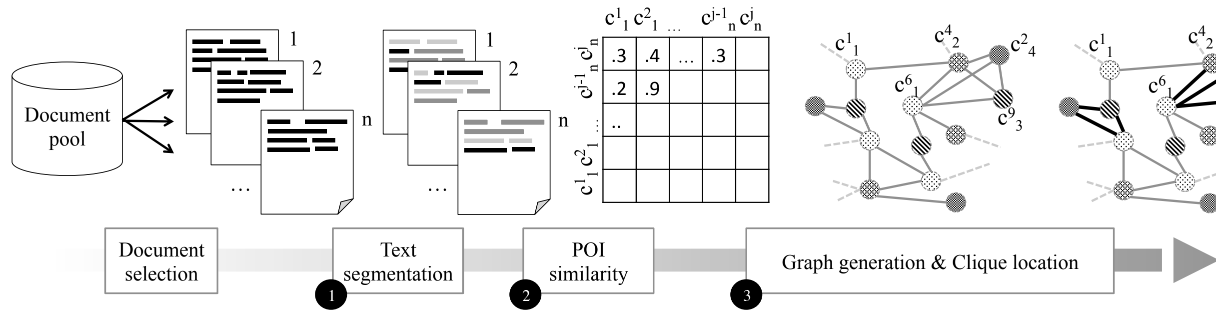

Figure 2: The proposed pipeline for locating cross-referenced pieces of information in news articles from various sources. A subset of articles pertaining to the same story is first selected from a pool. Their text is then segmented into meaningful pieces of information (called POIs), and the similarity between POIs is later computed. A graph is then generated from the POIs and their similarities. Cross-referenced POIs are those that form cliques in the graph.

approach locates POIs by concatenating neighboring sentences in a "meaningful" manner, that is by ensuring that important words are evenly distributed among POIs. To this end, for each of the sentences appearing in the set $\mathcal{A}_s$ we first compute its TF-IDF score, that is the average TF-IDF score of the words it contains. We then convert the segmentation problem into an optimization problem and employ a genetic algorithm (GA) that aims to minimize the standard deviation of the TF-IDF scores across POIs (although different optimization algorithms could have been used, or a multi-sentence segmentation version of TextTiling [9]).

POI similarity. In the second step, the pairwise similarity between POIs across documents in $\mathcal{A}_s$ is computed. For the purposes of the current system we decided to use the popular cosine TF-IDF similarity:

$$sim(v_i, v_j) = \frac{\sum_{w \in V}(v_i[w] \times v_j[w])}{\sqrt{\sum_{w \in V}(v_i[w])^2}\sqrt{\sum_{w \in V}(v_j[w])^2}} \tag{1}$$

where the v_i and v_j are vector representations of the compared POIs, and $v_i[w]$ corresponds to the product of the word's w frequency in the i^{th} POI and the IDF of w. Such a naive similarity function comes with both advantages and shortcomings. The method's focus on lexical features only is domain-agnostic, meaning it treats all texts equally without any pre-learned rules or associations. However, it fails to consider the similarity at the semantic level, and cannot handle polysemy. In the next iteration of the pipeline we will employ word and sentence embeddings approaches, e.g. [6, 11].

Graph Generation and clique location. Finally, in order to find cross-referenced POIs, we first convert the POI similarity matrix into a vertex-labeled weighted graph $G(V, E)$. Given a set of labels L corresponding to the document sources (e.g. CNN, Fox News), each vertex $v_i^l \in V$ corresponds to a POI with label $l \in L$. In addition, $E = \{(u_i^k, v_j^f) \in V \times V : k \neq f\}$ with the edge weight $w(u_i^k, v_j^f)$ corresponds to the POI similarity of the participating vertices.

Cross-referenced POIs correspond to the notion of cliques (or complete subgraphs) i.e., a subset of vertices that are all adjacent to each other. This correspondence is based on the idea that the information contained in a subset of vertices is likely to be credible if the POIs from different sources support each other's content.

Clique detection is an NP-complete problem and as such, various heuristic methods have been proposed. We use the method of Bron and Kerbosch [3] as provided by the `NetworkX Python` package. As a post-processing step, we remove any conflicting vertices, that is vertices that appear in two or more cliques (assuming that a piece of information can only be supported once per document). We perform this starting from the less important cliques, as measured using their TF-IDF score.

5 CREDIBILITY OF NEWS OUTLETS

We argue that locating cross-referenced information can be valuable for other purposes besides an interactive interface as described in Section 3. For example, it is intuitive to hypothesize that certain news outlets have a higher tendency than others to omit some credible information from their readers. Our proposed method can provide us with preliminary yet valuable insights concerning this hypothesis. As such, this section describes a case study which explores the relationship between corroborated and omitted information with regard to various news outlets.

Materials. We use the All-the-news dataset[4] which contains around 140.000 articles from January 2016 until July 2017 as published by fifteen major U.S. outlets including the New York Times, Breitbart, CNN and Fox News. For the sake of simplicity, we use only those articles corresponding to the whole year of 2016—85.405 articles in total. From those, we locate the sets of articles pertaining to the same story by employing our clique location approach - now on the article level. However, due to the dataset's size, prior to running the clique location we set a cutoff such that only cliques of articles with an average TF-IDF similarity > 0.6 are considered. This results in 203 cliques. We chose this cutoff after manual inspection as the resulting article sets were of good quality (wrt. pertaining to a single topic). At the next step, we locate the cross-referenced cliques of POIs for every clique of articles. In order to remove noisy

[4] https://www.kaggle.com/snapcrack/all-the-news, retrieved May 2018

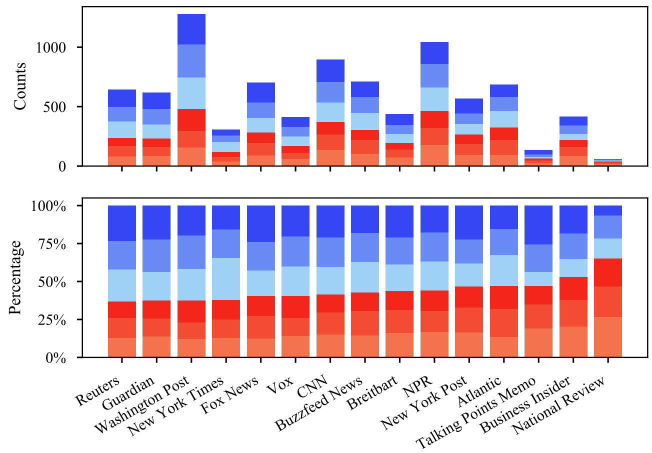

Figure 3: Raw amount (top) and percentage (bottom) of corroborated (blue) and omitted (orange) POIs per news outlet. Color shading (quantized to three bins) indicates the average TF-IDF similarity among POI cliques, i.e. the lighter the shade the more dissimilar the POIs.

cliques, based on manual examination we exclude those with an average TF-IDF similarity lower than 0.25.

Metric. We are interested in identifying the news outlets that provide their readers with high a *ratio of corroborated to omitted information.* This basic metric can act as an approximation of a news outlet's general credibility. Figure 3 presents the amount of corroborated and omitted POIs per outlet, shaded with different colors (blue and orange respectively) for the sake of visual inspection.

Results. A first general observation is that the amount of corroborated POIs is slightly larger than the omitted one; in average each outlet corroborates roughly 60% of the cross-referenced POIs in a clique of articles. However, it is the deviation from that trend that is most interesting. For example, Reuters and Guardian generally provide more corroborated than omitted information, close to a 2-to-1 ratio. At the same time, National Review exhibits the opposite behavior.

Discussion. It is interesting to note that independent projects aiming at exposing overtly biased news outlets, such as Media Bias/Fact Check[5], have evaluated Reuters and National Review is extremely objective and biased respectively; thus, supporting the results of our analysis. However, we should acknowledge that other outlets evaluated as biased, such as Vox or Breitbart, do not deviate from the general trend in our analysis. We suspect that such inconsistencies are related to the fact that we deal with articles only in cliques; thus, we do not consider editorials or independent articles in general that may shift the public opinion about an outlet's general bias. In addition, credibility (or bias in the case of Media Bias/Fact Check) is a multidimensional phenomenon and it is ill-advised to expect the ratio of corroborated to omitted information to model it completely.

[5] https://mediabiasfactcheck.com/, retrieved May 2018

6 CONCLUSIONS AND FUTURE WORK

Working towards aiding readers to make well-informed credibility judgments when consuming news online, this paper proposed a user interface design that is based on cross-referenced pieces of information in articles from different sources pertaining to the same story. We argue that our design fulfills the requirements for *transparency* and *explainability*: the reader can identify which news outlets have been used for cross-referencing information, and why certain pieces of information are considered more likely to be credible by the system.

This paper also presented a working pipeline for locating cross-referenced information that can find immediate application in real-life scenarios. We further showcased its potential by analyzing U.S. news outlets in terms of their corroborated/omitted information. An analysis of its output revealed some interesting findings regarding how much information is supported by other sources and how much information is potentially withheld from us, the readers.

However, further evaluation is required to provide insights regarding the accuracy of any method used within the proposed pipeline. Toward this aim, we plan to perform a user study in order evaluate the proposed interactive interface, and the efficiency of identifying corroborating or obfuscated information with regard to the readers' credibility judgment. An initial crowdsourcing-based user study yielded few insights as most crowd-workers did not engage to a meaningful degree with the interface (most—36 workers out of 43—spend their time simply waiting for the timer to be up to proceed to the post-study questionnaire); we are now exploring the deployment of our system in a social science Massive Open Online Course (MOOC) with potentially thousands of engaged learners using our tool.

Our future work will focus on improving the framework for locating cross-referenced information. Future implementations should consider employing more sophisticated short-text similarity methods such as those based on word and sentence embeddings [6, 11]. In addition, future implementations should consider the issue of document diversity: any system that deals with diverse information sources can only be perceived as credible if those sources promote multiple viewpoints equally [4]. Solving this issue would make the system invulnerable to possible scams, such as multiple articles promoting the same false story.

Finally, we aim to apply our interface design to other related tasks in need of explainability, such as news recommendation, or text plagiarism and copyright infringement.

ACKNOWLEDGMENTS

This research has been supported by NWA Startimpuls VWData, Delft Data Science and NWO project SearchX (639.022.722).

REFERENCES

[1] James Allan. 2002. Introduction to topic detection and tracking. In *Topic detection and tracking*. Springer, 1–16.
[2] Engin Bozdag and Jeroen van den Hoven. 2015. Breaking the filter bubble: democracy and design. *Ethics and Information Technology* 17, 4 (2015), 249–265.
[3] Coen Bron and Joep Kerbosch. 1973. Algorithm 457: finding all cliques of an undirected graph. *Commun. ACM* 16, 9 (1973), 575–577.
[4] Sylvie Cazalens, Philippe Lamarre, Julien Leblay, Ioana Manolescu, and Xavier Tannier. 2018. A Content Management Perspective on Fact-Checking. In *Proceedings of the Web Conference 2018-alternate paper tracks" Journalism, Misinformation and Fact Checking"*. 1–10.

[5] Niall J Conroy, Victoria L Rubin, and Yimin Chen. 2015. Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology* 52, 1 (2015), 1–4.
[6] Cedric De Boom, Steven Van Canneyt, Steven Bohez, Thomas Demeester, and Bart Dhoedt. 2015. Learning semantic similarity for very short texts. In *Proceedings of the Data Mining Workshop*. 1229–1234.
[7] Mitja Trampuš Flavio Fuart, Daniele Pighin, Tadej Štajner, Jan Berčič, Blaž Novak, Delia Rusu, Luka Stopar, and Marko Grobelnik. 2015. DiversiNews: Surfacing Diversity in Online News. *AI Magazine* 36, 4 (2015).
[8] Felix Hamborg, Norman Meuschke, and Bela Gipp. 2017. Matrix-based news aggregation: exploring different news perspectives. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*. IEEE, 1–10.
[9] Marti A Hearst. 1997. TextTiling: Segmenting text into multi-paragraph subtopic passages. *Computational linguistics* 23, 1 (1997), 33–64.
[10] Benjamin D Horne, William Dron, Sara Khedr, and Sibel Adali. 2018. Assessing the News Landscape: A Multi-Module Toolkit for Evaluating the Credibility of News. In *Companion of the The Web Conference 2018 on The Web Conference 2018*. International World Wide Web Conferences Steering Committee, 235–238.
[11] Tom Kenter and Maarten De Rijke. 2015. Short text similarity with word embeddings. In *Proceedings of the international Conference on Information and Knowledge Management*. ACM, 1411–1420.
[12] Miriam J Metzger. 2007. Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the Association for Information Science and Technology* 58, 13 (2007), 2078–2091.
[13] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. 2009. NewsCube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the Conference on Human Factors in Computing Systems*. ACM, 443–452.
[14] VG Vydiswaran, ChengXiang Zhai, and Dan Roth. 2011. Content-driven trust propagation framework. In *Proceedings of the international conference on Knowledge Discovery and Data Mining*. ACM, 974–982.
[15] C Nadine Wathen and Jacquelyn Burkell. 2002. Believe it or not: Factors influencing credibility on the Web. *Journal of the Association for Information Science and Technology* 53, 2 (2002), 134–144.
[16] Xiaoxin Yin, Jiawei Han, and S Yu Philip. 2008. Truth discovery with multiple conflicting information providers on the web. *IEEE Transactions on Knowledge and Data Engineering* 20, 6 (2008), 796–808.